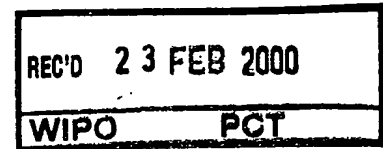


**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)



DE 99/3922

## **Bescheinigung**

Die Siemens Aktiengesellschaft in München/Deutschland hat eine Patentanmeldung  
unter der Bezeichnung

"Controlling- und Marketingverfahren für die Nutzung  
des Internet/Intranet"

am 11. Dezember 1998 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprüng-  
lichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig das Symbol  
G 06 F 17/60 der Internationalen Patentklassifikation erhalten.

München, den 14. Januar 2000

**Deutsches Patent- und Markenamt**

**Der Präsident**

Im Auftrag

Jerofsky

Aktenzeichen: 198 57 336.7

**THIS PAGE BLANK (USPIC)**

## Beschreibung

### Controlling- und Marketingverfahren für die Nutzung des Internet/Intranet

5

Der Anmeldungsgegenstand betrifft ein Verfahren zur Erfassung des Nutzungsverhaltens eines Teilnehmers im Internet/Intranet umfassend die Merkmale des Anspruchs 1.

- 10 Nicht vertragsgemäße Nutzung des Internetzugangs (z.B. von Privatpersonen bei flat rate Tarifen eines Internet Service Provider oder von Angestellten einer Firma) kann eine Größenordnung erreichen, die die Geschäftsbasis eines Internet Service Provider oder einer Firma gefährdet. Derzeit ist es
- 15 technisch sehr schwierig, solche Nutzungsfälle so frühzeitig zu erkennen, daß rechtzeitig reagiert werden kann.

- Privat- und Geschäftskunden oder auch Angestellte einer Firma nutzen das Internet/Intranet auf unterschiedliche Art und
- 20 Weise, z.B. zum gelegentlichen Surfen, zur Übertragung von Datenraten (in unterschiedlicher Größe), zum Spielen, etc. Gezielte Marketingaktionen für die unterschiedlichen Nutzungskreise seitens des Internet Service Provider, Erkennen von Markttrends oder auch das Bestimmen von Kosteneinsparungspotentialen einer Firma (z.B. bei Corporate Networks) sind bisher nur recht schwer möglich, da eine genaue Zuordnung eines Internet/Intranet-Nutzers zu diversen Verhaltens-

- 30 Es gibt derzeit diverse Tools zur Erkennung von Betrugsfällen in der Telekommunikation. Solche Tools basieren auf verschiedenen Techniken, wie den regelbasierten Ansatz oder neuronale Netze etc. Ausgewertet werden mit diesen Techniken CDR (Call Detail Records) oder Signalisierungsdaten des Zeichengabesystems CCS7.
- 35

Es ist ein Werkzeug unter der Bezeichnung "HP OpenView Smart Internet Suite; Smart Internet Usage" angekündigt worden, das nutzungs-spezifische Internet-Daten sammelt, korreliert und verdichtet sowie auf diese Daten eine Retrieval-Funktion (data mining) anbietet. Die technische Realisierung und der genaue Umfang der Retrieval-Funktion sind im Detail nicht bekannt geworden.

Dem Anmeldungsgegenstand liegt das Problem zugrunde, ein Verfahren anzugeben, das gegenüber herkömmlichen Verfahren eine Erhöhung der Signifikanz der Aussage mit verringerter Fehlerquote aufweist.

Das Problem wird durch ein Verfahren mit den Merkmalen des Anspruchs 1 gelöst.

Durch dieses Werkzeug können sowohl frühzeitig Verdachtsmomente auf nicht-spezifische Nutzungen (z.B. nicht vertragsgemäße Nutzung des Internet / Intranet) gesammelt werden (Stichwort: Betrug), als auch Ergebnisse erzielt werden, die für Marketingzwecke, für eine Markttrenderkennung, für schnelle Reaktionen hinsichtlich eines notwendigen Netzausbaus des Internet / Intranet oder auch zur Einsparung von Kosten bei der Nutzung des Internet / Intranet genutzt werden können.

Durch die Erfindung erhalten Internet Service Provider und Firmen extrem gute Aussagen zu - Art der Nutzung des Internet / Intranet (insbesondere auch zu einer nicht vertragsgemäßen Nutzung) - Markttrends (insbesondere auch zu abrupten Verhaltensänderungen hinsichtlich der Nutzung von Internet / Intranet) - Marketing- zu einem notwendigen Netzausbau etc. Insbesondere werden durch diese Erfindung die Nachteile der auf dem Markt befindlichen Data Mining Tools beseitigt. Gemäß einer besonderen Ausgestaltung des Anmeldungsgegenstandes werden die Ergebnisse der einzelnen Methoden zu einer si-

gnifikanten Aussage mit äußerst geringer Fehlerquote verdichtet (kombiniert, verknüpft).

5 Vorteilhafte Weiterbildungen des Anmeldungsgegenstandes sind in den Unteransprüchen angegeben.

Der Anmeldungsgegenstand wird im folgenden als Ausführungsbeispiel in einem zum Verständnis erforderlichen Umfang anhand von Figuren näher erläutert. Dabei zeigen:

10 Fig 1 eine bekannt gewordene Ablauffolge,  
Fig 2 eine prinzipielle Blockdarstellung von Elementen und deren Zusammenwirken beim Anmeldungsgegenstand und  
Fig 3 ein Anwendungsbeispiel für die Modellierung einer Verhaltenskategorie im kausalen Netz.

15 In den Figuren bezeichnen gleiche Bezeichnungen gleiche Elemente.

Das bekannt gewordene Werkzeug arbeitet - soweit erkennbar -  
20 wie in Fig 1 dargestellt folgendermaßen: In einem 1. Schritt werden die Internet-Daten Idat in einer Vorverarbeitung PP (für: preprocessing) nach festen Regeln zu Internet Data Records (IDR) korreliert und verdichtet; in einem 2. Schritt sind die Internet Data Records mit vom Operator eingegebenen Regeln (z. B. in Form von Select Statements) in einem Bewertungsvorgang RETR (für: retrieval, Data mining) als Ergebnis OUT ausgebbar. Dabei macht es sich nachteilig bemerkbar, daß  
- Eine Änderung der IDR eine Änderung des Preprocessor bedingt - die Regeln sich nur auf die IDR (insbesondere auch  
30 IDR-Struktur) beziehen dürfen.- die IDR statisch sind, wodurch Informationen verlorengehen, - rein regel-basierte Systeme nicht lernfähig sind, - rein regel-basierte Systeme keine "Ausnahmen von der Regel" kennen (damit große Anzahl von Fehlaussagen)- rein regel-basierte Systeme keine Grenzbe-  
35 reiche kennen.

Bei dem in Fig 2 dargestellten anmeldungsgemäßen Verfahren können die Internet-Daten Idat einer Regel-basierten Vorver-

arbeitung RBPP (für: Regel-basierter Preprocessor) unterzogen werden, wobei die Internet-Daten korreliert und verdichtet werden. Die Internet-Daten können, gegebenenfalls in vorverarbeiteter Form, als Zwischenergebnis einer Zwischenspeicherung INTM(für: Interim Memory) unterzogen werden. Die Internet-Daten werden, gegebenenfalls in vorverarbeiteter Form und gegebenenfalls nach Zwischenspeicherung, einer Einrichtung Methodenansatz MA zugeführt, die einen regelbasierten Ansatz RBA, ein neuronales Netz mit überwachtem Training NNUE, eine dichte-basierte Profilmodellierung DBPM und ein kausal neuronales Netz KNN, im folgenden kausales Netz genannt, aufweist. Die Einrichtung Methodenansatz MA arbeitet, wie durch zwei doppeltgerichtete Pfeile bezeichnet, mit einer Regelbasis RB, in der die Regeln abgespeichert sind, einer Datenbasis MO/TR, in der die Modellierungs-/Trainingsdaten abgespeichert sind, und einer Datenbasis HIST, in der die Auswertergebnisse des aktuellen und vorangegangener Beobachtungszeiträume abgespeichert sind, zusammen. Die von der Einrichtung Methodenansatz MA ausgegebenen bzw. in der Datenbasis HIST gespeicherten Zwischenergebnisse können in einer Einrichtung COMB (für: Combination) einer Bewertung unterzogen werden und werden als Ergebnis OUT ausgegeben.

Das anmeldungsgemäße Verfahren umfaßt eine Kombination von vier unterschiedlichen Methodenansätzen, den regel-basierten Ansatz und drei Methodenansätze der Neuroinformatik (neuronales Netz mit überwachtem Training, dichte-basierte Modellierung und kausales Netz). Die Regelformulierung bzw. die Modellierung mit den Methoden der Neuroinformatik werden auf der Basis von Daten durchgeführt, die der Internet Service Provider oder eine Firma speichert: RADIUS Accounting Daten (werden in der Regel gespeichert), TCP Dump Protokoll Daten (Speicherung bei Bedarf, Umfang variabel), SNMP(Simple Network Management Protocol)-Daten (Speicherung bei Bedarf) etc. Das entstehende Modell stellt ein Controlling- und Marketingwerkzeug dar.

Das anmeldungsgemäße Verfahren ist vorgesehen zur Anwendung auf Internet / Intranet-Daten, die der Internet Service Provider oder eine Firma speichert. Solche Daten sind u.a. die RADIUS Accounting Daten, die TCP Dump Daten und die SNMP Daten. Das Verfahren kann beliebige weitere Internet / Intranet-Daten bearbeiten.

Die RADIUS Accounting Daten sind durch Daten gegeben, wie sie in der IETF Spezifikation RFC 2139 beschrieben sind. Eine reale Implementierung ist z.B. in Livingston Enterprises Inc., Radius dictionary, V1.6, 1997 beschrieben.

Die TCP Dump Daten sind durch Daten gegeben, wie sie in UNIX man-pages 'tcpdump - dump traffic on a network' beschrieben sind.

Die SNMP Daten sind durch Daten gegeben, wie sie in den verschiedenen RFCs der IETF beschrieben sind. Eine tatsächliche Implementierung ist z.B. in Livingston Enterprises Inc., Configuring SNMP, Manual Portmaster 3 beschrieben.

Um die Verarbeitung der Daten zu beschleunigen, kann optional ein regel-basierter Preprocessor eingesetzt werden. Der Preprocessor hat die Aufgabe, die Internet / Intranet-Daten so zu korrelieren und zu verdichten, daß als Ergebnis Datensätze mit den im eigentlichen Verfahren benötigten Attributwerten geliefert werden.

Grundsätzlich kann ein Preprocessor verwendet werden, wie er bei dem bekannt gewordenen Lösungsvorschlag vorgesehen ist. Dies setzt aber voraus, daß die IDR eine Obermenge der vom Verfahren benötigten Attributwerte beinhalten.

In einer bevorzugten Ausführungsform des Anmeldungsgegenstandes wird ein regel-basierter Preprocessor verwendet. Die Regeln steuern dabei die Korrelation und die Verdichtung der Internet / Intranet-Daten.

Wird im eigentlichen Verfahren ein neues charakteristisches Attribut hinzugefügt oder fällt ein charakteristisches Attribut weg, dann können die Auswahlregeln des Preprocessor einfach (automatisch) angepaßt werden. Eine automatische Anpassung der Auswahlregeln kann dabei, wie in der Fig 2 mit ADAP (für: Adaption) bezeichnet, über Notifications (freilaufende Meldungen) an den Preprocessor gesteuert werden.

- 10 Das nachfolgend beschriebene Verfahren kann damit
- direkt auf den Internet / Intranet-Daten aufsetzen,
  - auf Ergebnisse eines beliebigen Preprocessors (wie z.B. HP(Hewlett Packard) IPR) aufsetzen
  - auf Ergebnisse (in Fig 2 mit INTM bezeichnet) eines speziellen regel-basierten Preprocessor aufsetzen.
- 15

Das eigentliche Verfahren teilt sich in vier Verfahren. Jedes Verfahren verwendet dabei einen anderen Methodenansatz. Die vier unterschiedliche Methodenansätze sind:

- 20
- das regelbasierte Verfahren,
  - das neuronale Netz - überwachtes Lernen
  - die dichte-basierte Profilmodellierung
  - das kausale Netz.

- 25 Beim regelbasierten Verfahren lassen sich typische benutzer-spezifische Verhaltenskategorien mit Hilfe von Regeln modellieren. Durch eine Verhaltenskategorie wird das Verhalten klassifiziert, z.B. können die Verhaltenskategorien "Privatnutzung Student", "Privatnutzung Angestellter", "Privatnutzung Freiberufler", "Nutzung kleiner Geschäftsbetrieb", "Nutzung großer Geschäftsbetrieb", "Spieler", "Internet / Intranet-Süchtiger", "Nutzer mit hohem Mailvolumen" etc. durch ihre charakteristischen Eigenschaften in Form von Regeln ausgedrückt werden. Die Regeln werden auf alle Internet / Intranet-Daten oder einen Teil dieser Daten (z.B. das Ergebnis eines Preprocessing) angewendet. Das Ergebnis des Verfahrens ist, daß jeder Nutzer nach einem Beobachtungs-zeitraum  $t$  kei-
- 30
- 35



ner, einer oder mehrerer Verhaltenskategorien zugeordnet werden kann. Der Beobachtungszeitraum kann dabei abhängig von der Verhaltenskategorie und dem gewünschten Zweck der Beobachtung variieren. So wird er zur Betrugsentdeckung (als Teil  
5 des Controlling) sehr klein gewählt werden (z.B.  $t = 1$  Tag), dagegen zur Erzielung von Marketingaussagen einen Zeitraum von mehreren Wochen umfassen (z.B.  $t = 4$  Wochen). Speichert man die Ergebnisse eines jeden Beobachtungszeitraums  $t(i)$  nut-  
10 zerspezifisch in der Datenbasis HIST, so lassen sich durch Vergleich der Einzelergebnisse  $t(i)$  recht einfach Änderungen im Benutzerverhalten feststellen (z.B. am Anfang entsprach die Nutzung eines Nutzers der Verhaltenskategorie "Privatnutzung Student", nunmehr ist sie eher der Verhaltenskategorie "Nutzung kleiner Geschäftsbetrieb" zuzuordnen).

15

Das Ziel ist es, Regeln für jede Verhaltenskategorie zu formulieren. Die Regeln werden mit Hilfe von logischen Ausdrücken beschrieben, in denen als Variablen die Felder (Attribute) der verschiedenen Datensätze verwendet werden, z.B.

20 Nutzung "Privatvertrag Angestellter" ::= für alle Datensätze des Beobachtungszeitraums gilt: Nutzungszeit Montag bis Freitag zwischen 17 Uhr und 24 Uhr und Nutzungszeit an Wochenenden von 0 Uhr bis 24 Uhr und Datenübertragungsrate  $< 2$  Megabyte pro Nutzung und max. Nutzungszeit = 2 Stunden)

Nutzung "Privatvertrag Freiberufler" ::= es gibt einen Datensatz des Beobachtungszeitraums, für den gilt: Nicht Nutzung "Privatvertrag Angestellter" und für alle Datensätze des Beobachtungszeitraums gilt: Datenübertragungsrate  $< 10$  Megabyte pro Tag und max. Nutzungszeit = 8 Stunden

30 Grundsätzlich können sich die Regeln auf einen oder mehrere Datensätze (auch verschiedener Dateien) beziehen.

In der Anwendungsphase des regelbasierten Verfahrens werden für die angegebenen Daten zum Zeitpunkt  $t$  alle ausgewählten  
35 Regeln überprüft. Die Ergebnisse werden zunächst benutzerspezifisch in der Datenbasis HIST protokolliert.

Stärke des regelbasierten Verfahrens:

- Klassifizierung des Benutzerverhaltens in Form von Verhaltenskategorien
- Hieraus lassen sich einfach Trends, Marketingaussagen, Vertragsverletzungen etc. ableiten

- Beim überwachten Ansatz wird ein Neuronales Netz mit einem Satz von Beispielen trainiert. Die Voraussetzung für das Training ist, daß zu jedem Beispiel der zugehörige Zielwert gegeben ist, d.h. es muß zum Zeitpunkt des Trainings bekannt sein, ob z.B. für das betrachtete Beispiel eine festgelegte Nutzung vorlag oder nicht (festgelegte Nutzungen können z.B. sein "Vertragsverletzung Privatvertrag Angestellter", "Schwerpunktnutzung Surfen", "Schwerpunktnutzung Spieler" etc.). Dabei müssen die zu untersuchenden Zielwerte und die für das Beispiel charakteristischen Attribute vorgegeben werden. Die charakteristischen Attribute bestimmen das Verhalten eines Nutzers. Das Verhalten wiederum ist abhängig von bestimmten Attributwerten (den Daten selber).
- Charakteristische Attribute können z.B. sein:
- mittlere Nutzungszeit des Nutzers an einem Tag über einen Beobachtungszeitraum (z.B. vier Wochen)
  - Streuung der Nutzungszeit des Nutzers
  - maximale Nutzungszeit
  - minimale Nutzungszeit
  - mittlere Übertragungsrate des Nutzers an einem Tag über einen Beobachtungszeitraum (z.B. vier Wochen)
  - Streuung der Übertragungsrate des Nutzers
  - maximale Rate
  - minimale Rate
  - mittlere Nutzungsdauer des Nutzers von speziellen Internet/Intranet Services über einen Beobachtungszeitraum (z.B. vier Wochen)
  - Streuung der Nutzungsdauer des Nutzers
  - maximale Nutzungsdauer
  - minimale Nutzungsdauer
- etc.

In der Trainingsphase (Vorabschritte) des Neuronalen Netzes ist es Ziel, ein Modell zu erstellen, das für einen Nutzer anhand des gegebenen Beispiels entscheidet, ob eine Nutzung des Internet/Intranetz Zugangs hinsichtlich eines oder mehrerer definierter Zielwerte stattfindet oder nicht. Die Modellerstellung erfolgt durch das überwachte Training, dessen Grundlagen in Rumelhart, D.E., Hinton, G.E. und Williams, R.J. Learning internal representation by error backpropagation, In Parallel Distributed Processing, S. 318-362, Cambridge, MA, MIT Press, 1986) im einzelnen beschrieben sind.

In der Trainingsphase werden folgende Schritte durchgeführt: Jedem Nutzer wird ein Verhaltensmuster in Form von Attributen zugeordnet, das ein gewisses Profil über einen längeren Zeitraum beschreibt. Die Attribute charakterisieren dabei die Nutzung hinsichtlich eines definierten Zielwertes. Der dem Verhaltensmuster zugrundeliegende Zeitraum sollte nicht kürzer als vier Wochen sein und vor dem Zeitpunkt liegen, in dem die Methode für o.g. Zweck angewendet wird.

Das neuronale Netz wird anhand von Trainingsdaten auf die Nutzung hinsichtlich der definierten Zielwerte trainiert. Bei den Trainingsdaten ist bekannt, ob die Nutzung einem bestimmten Zielwert zugeordnet werden kann oder nicht.

In der Anwendungsphase des neuronalen Netzes, die nach Abschluß der Trainingsphase beginnt, werden kontinuierlich folgende Schritte durchgeführt:

30

Das Neuronale Netz entscheidet auf Basis der Beispiele, ob die Nutzung einem bestimmten Zielwert zugeordnet werden kann oder nicht. Diese Entscheidung wird nutzer-spezifisch als Ergebnis des Beobachtungszeitraumes in der Datenbasis HIST protokolliert.

35

Das neuronale Netz kann bei Bedarf mit neuen Zielwerten hinsichtlich seiner Nutzung (z.B. bis dato noch nicht bekannten Fällen von Vertragsverletzungen) trainiert werden.

- 5 Diese Methode anwendbar, wenn der Nutzer Bestandteil der Daten ist.

Stärke dieser Methode:

- Einfache Zuordnung von Ausnahmen
- 10 - Berücksichtigung von Ausnahmen im Ergebnis
- Lernfähigkeit

Bei der dichte- basierten Profilmodellierung handelt es sich um eine probabilistische Modellierung des Verhaltens zu jedem  
15 Nutzer (probabilistische Profilmodellierung), d.h. es wird für jeden Nutzer ein Modell anhand der zu diesem Nutzer zugehörigen Beispielen erstellt. Diese Beispiele bestehen aus charakteristischen Attributen und aus bestimmten Attributwerten, die die Nutzung des Internet/Intranet hinsichtlich eines  
20 oder mehrerer Zielwerte beschreiben. Beispiele für charakteristische Attribute sind im vorhergehenden Abschnitt beschrieben.

In der Trainingsphase der dichte-basierten Profilmodellierung werden folgende Schritte durchgeführt: Jedem Nutzer wird ein  
25 Satz von Beispielen zugeordnet, der die Verhaltensweise des Nutzers über einen längeren Zeitraum beschreibt. Der dem Verhaltensmuster zugrundeliegende Zeitraum sollte nicht kürzer als vier Wochen sein und vor dem Zeitpunkt liegen, in dem die Methode für Kontroll- und Marketingzwecke angewendet wird.  
30 Für jeden Nutzer wird ein probabilistisches Profil erstellt. Diese Erstellung erfolgt durch die Dichteschätzung mit dem EM-Algorithmus. Die genaue Beschreibung ist in Chris Bishop, Neural Networks in Pattern Recognition, Oxford Press, 1996 enthalten. Nach Abschluß der Trainingsphase, beginnt die Anwendungsphase der dichte-basierten Profilmodellierung, in der  
35 kontinuierlich folgende Schritte durchgeführt werden: Die Daten beispielsweise eines Tages werden hinsichtlich der für

die probabilistische Profilmodellierung bestimmten Dateninhalte analysiert (ein neues Beispiel wird erzeugt). Das dichte-basierte Profilmodell gibt einen Wert aus, der eine Wahrscheinlichkeit für die Nutzung des Internet/Intranet der betrachteten Entität hinsichtlich der definieren Zielwerte darstellt. Dieser Wert wird protokolliert. Weicht dieser Wert gegenüber den vorangegangenen Werten über einen vorgegebenen Schwellwert hinaus ab, dann erfolgt ein Hinweis, daß das Ergebnis in jedem Fall angezeigt werden soll. Mit dieser Methode kann sehr einfach festgestellt werden, wenn sich die Nutzung des Internet/Intranet plötzlich ändert. Mit dem aktuellen Beispiel wird das Profilmodell nachadaptiert. Diese Methode ist anwendbar, wenn der Nutzer Bestandteil der Daten ist. Stärke dieser Methode: - Erkennung einer abrupten Änderung des Benutzerverhaltens - Lernfähigkeit

Basis für die Methode des kausalen Netz' ist die Modellierung typischer Verhaltensszenarien in Form von kausalen Abhängigkeiten und Wahrscheinlichkeiten bestimmter Dateninhalte wie es in dem Beispiel "Privatnutzung Angestellter" in Fig 3 dargestellt ist. Einer Privatnutzung Angestellter PA sind eine bestimmte Nutzungszeit UC (für: UseClock), eine bestimmte Nutzungsdauer UT (für: UseTime) und eine bestimmte Übertragungsrate RATE zugeordnet. Die Wochentage haben, je nachdem ob es sich um einen Werktag WD (für: working day) oder um ein Wochenende WE (für: week end) handelt, Einfluß auf die Höhe der Nutzungszeit, die Nutzungsdauer und die Übertragungsrate. Die kausalen Abhängigkeiten basieren auf der Auswertung bekannter Fälle. Sie müssen nicht spezifischen Nutzern zugeordnet sein. In der Phase der Modellierung des kausalen Netzes werden folgende Schritte durchgeführt: Für alle Daten werden die kausalen Abhängigkeiten hinsichtlich der Daten-Inhalte formuliert. An den Stellen, wo die kausalen Abhängigkeiten existieren, werden entsprechende Wahrscheinlichkeiten zugeordnet. In der Modellierungsphase ist das Domänenwissen des Fachexperten notwendig. Die Grundlage über das kausale Netz ist in Finn V. Jensen, An Introduction to Bayesian Networks, UCL Press 1996 beschrieben. In der Anwendungsphase des kaus-

len Netzes werden kontinuierlich folgende Schritte durchgeführt: Die Datensätze der zu untersuchenden Daten werden kontinuierlich auf die formulierten kausalen Abhängigkeiten hin untersucht. Für jeden Nutzer bzw. jedes Ereignis wird entschieden, mit welcher Wahrscheinlichkeit eine spezifische Nutzung hinsichtlich des definierten Szenarios vorliegt. Diese Entscheidung wird als Ergebnis des Beobachtungszeitraumes in der Datenbasis HIST protokolliert. Ist der Nutzer Bestandteil der Daten, dann werden die Ergebnisse nutzer-spezifisch protokolliert. Die Wahrscheinlichkeiten hinter den kausalen Abhängigkeiten können nachadaptiert werden. Die kausalen Abhängigkeiten von neuen, bis dato noch nicht bekannten Kategorien werden bei Bedarf den existierenden kausalen Abhängigkeiten hinzugefügt. Diese Methode läßt sich auch dann anwenden, wenn der Nutzer nicht Bestandteil des Datensatzes ist. In diesem Fall kann aber eine Kategorie keinem spezifischen Nutzer zugeordnet werden. Stärke der Methode des kausalen Netzes: - Zuordnung von Grenzbereichen in Verhaltenskategorien - Erkennen von Grenzbereichen in Verhaltenskategorien - Lernfähigkeit Grundsätzlich ist es möglich, die Einzelergebnisse der einzelnen Verfahren auszugeben. In weiterer Ausgestaltung des Anmeldungsgegenstandes werden die Einzelergebnisse der einzelnen Verfahren zu einem Gesamtergebnis verdichtet. Diese Verdichtung bezieht die Einzelergebnisse der verschiedenen Verfahren ein. Die Einzelergebnisse können sowohl aus dem aktuellen als auch aus vergangenen Beobachtungszeiträumen stammen. Jedes der beschriebenen Verfahren hat spezifische Stärken. Dies kann bei einer Verdichtung ausgenutzt werden, wie im folgenden Beispiel beschrieben: Ein Nutzer x wird z.B. durch das regel-basierte Verfahren eindeutig der Verhaltenskategorie "Privatnutzung Freiberufler" zugeordnet, wenn er an einem Tag das Internet mehr als 2 Stunden genutzt hat (bezogen auf die oben dargestellte Beispielregel). Das kausale Netz sieht diesen Nutzer x aber vielmehr im Verhaltensszenario "Privatnutzung Angestellter", da er z.B. in mehr als 90% der Datensätze die Nutzungszeit von weniger als 2 Stunden eingehalten hat. Diese Erkenntnisse könnten dann im Gesamter-

gebnis dergestalt angezeigt werden, daß es sich bei der Nutzung des Internet / Intranet durch den Nutzer x mit einigen geringfügigen Ausnahmen um eine "Privatnutzung Angestellter" handelt. Ein weiteres Beispiel für eine Verdichtung ist die

- 5 Trenderkennung durch Auswertung von Ergebnissen verschiedener Beobachtungszeiträume.

## Patentansprüche

1. Verfahren zur Erfassung des Nutzungsverhaltens eines Teilnehmers im Internet/Intranet, demzufolge

5 - die Teilnehmerdaten über einen festgelegten Zeitraum protokolliert werden

- die protokollierten Teilnehmerdaten nach mindestens einem der folgenden Verfahren ausgewertet werden:

- Regel-basierter Ansatz,
- 10 • neuronales Netz, überwacht trainiert,
- dichte basierte Profilmodellierung,
- kausales Netz

derart, daß eine Zuordnung eines Teilnehmers nach Maßgabe seines Teilnehmer-Verhaltens zu einer Kategorie von Nutzungsverhalten gegeben ist.

15

2. Verfahren nach Anspruch 1,  
dadurch gekennzeichnet,

20 daß die Teilnehmerdaten durch mindestens eine der folgenden Kategorien von Daten gegeben sind: RADIUS-Daten, SNMP-Daten, TCP-Dump-Protokoll-Daten.

3. Verfahren nach einem der vorstehenden Ansprüche,  
dadurch gekennzeichnet,

25 daß die Teilnehmerdaten vor der Auswertung einer Regel-basierten Vorbehandlung unterzogen werden.

4. Verfahren nach Anspruch 3,  
dadurch gekennzeichnet,

30 daß die einer Vorbehandlung unterzogenen Teilnehmerdaten vor der Auswertung zwischengespeichert werden.

5. Verfahren nach einem der Ansprüche 3 oder 4,  
dadurch gekennzeichnet,

35 daß nach Maßgabe des Auswerteergebnisses mindestens eines der Verfahren eine automatische Anpassung der Auswahlregeln für die Vorbehandlung der Teilnehmerdaten erfolgt.



6. Verfahren nach einem der vorstehenden Ansprüche,  
dadurch gekennzeichnet,  
daß die Ergebnisse, insbesondere die aktuellen Ergebnisse,  
5 der Auswertungen der einzelnen Verfahren zu einem Endergebnis  
verknüpft werden.

7. Verfahren nach einem der vorstehenden Ansprüche,  
dadurch gekennzeichnet,  
10 daß die Ergebnisse der Auswertungen der einzelnen Verfahren  
eines aktuellen Beobachtungszeitraumes und eines vergangenen  
Beobachtungszeitraumes zu einem Endergebnis verknüpft wer-  
den.

## Zusammenfassung

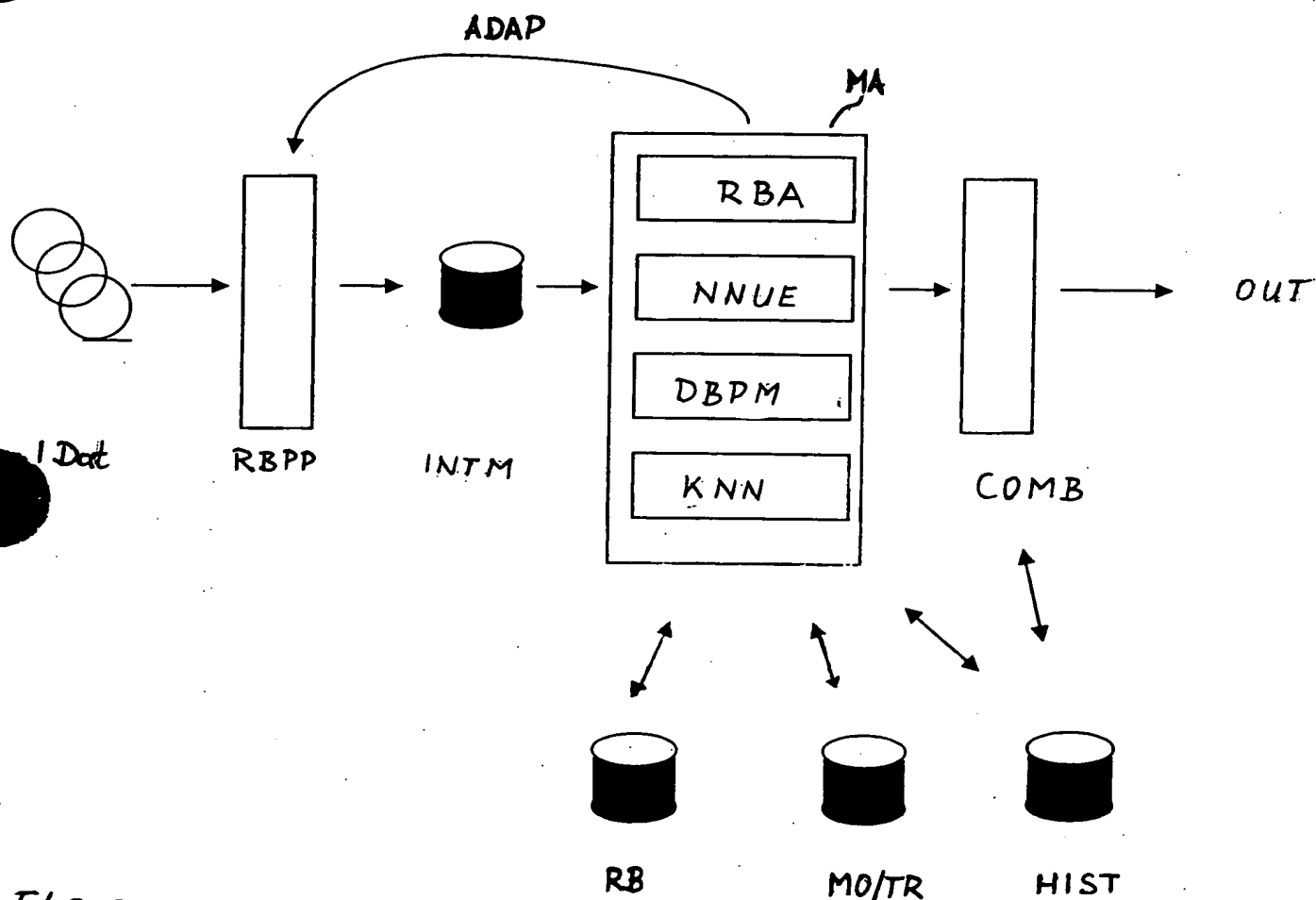
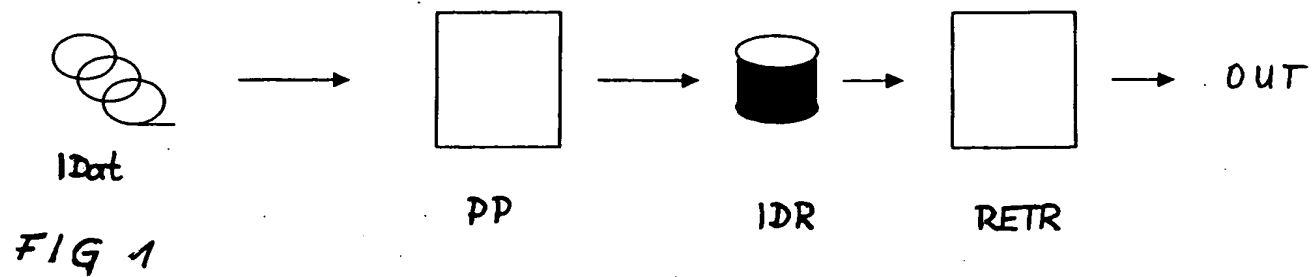
Controlling- und Marketingverfahren für die Nutzung des Internet/Intranet

5

Auf die Daten eines Teilnehmers im Internet/Intranet werden verschiedene Methoden wie der regelbasierte Ansatz, ein neuronales Netz mit überwachtem Training, eine dichte-basierte Profilmmodellierung und ein kausales Netz angewendet und die  
10 Ergebnisse daraus gegebenenfalls kombiniert, um für das Nutzungsverhalten eine Erhöhung der Signifikanz der Aussage mit verringerter Fehlerquote zu erzielen, wodurch eine frühzeitige Erkennung von Unregelmäßigkeiten im Nutzungsverhalten, Aussagen über die Notwendigkeit eines Netzausbaus sowie si-  
15 gnifikante Aussagen für Marketingzwecke und über Kosten gegeben sind.

Fig 2

20



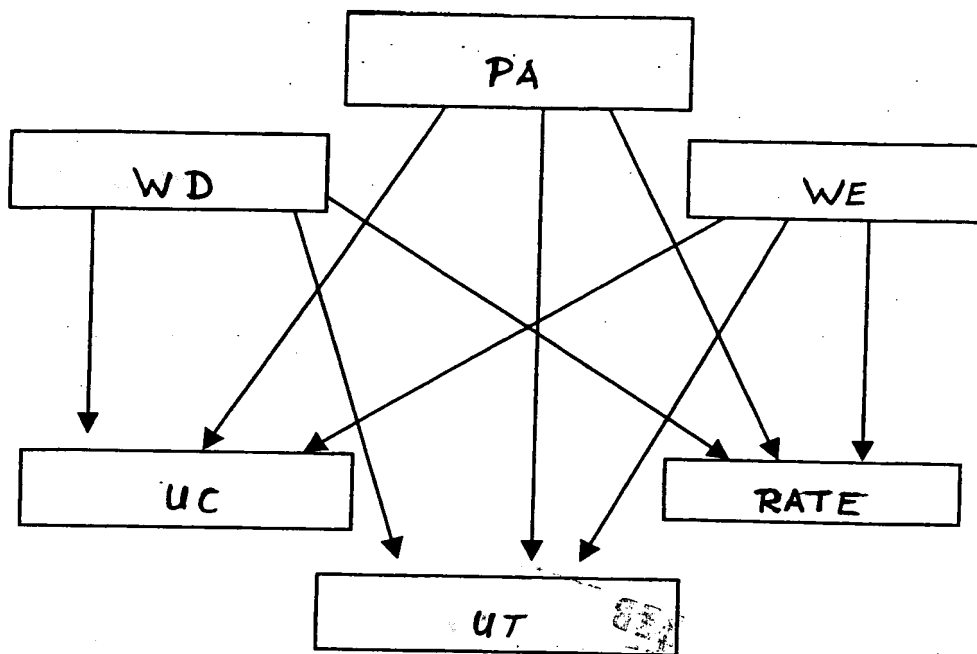


FIG 3